Fisheries and Oceans Canada | Pêches et Océans Canada
Ecosystems and Oceans Science | Sciences des écosystèmes et des océans

Canadian Science Advisory Secretariat (CSAS)

Research Document 2016/064

Gulf Region

Abundance indices of Atlantic herring (*Clupea harengus*) from the southern Gulf of St. Lawrence based on the September multispecies bottom trawl survey

Tobie J. Surette

Fisheries and Oceans Canada
Gulf Fisheries Centre
343 University Avenue, P.O. Box 5030
Moncton, New Brunswick E1C 9B6

October 2016

Canada

Foreword

This series documents the scientific basis for the evaluation of aquatic resources and ecosystems in Canada. As such, it addresses the issues of the day in the time frames required and the documents it contains are not intended as definitive statements on the subjects addressed but rather as progress reports on ongoing investigations.

Research documents are produced in the official language in which they are provided to the Secretariat.

Published by:

Fisheries and Oceans Canada
Canadian Science Advisory Secretariat
200 Kent Street
Ottawa ON K1A 0E6

http://www.dfo-mpo.gc.ca/csas-sccs/
csas-sccs@dfo-mpo.gc.ca

© Her Majesty the Queen in Right of Canada, 2016
ISSN 1919-5044

Correct citation for this publication:

Surette, T.J. 2016. Abundance indices of Atlantic herring (*Clupea harengus*) from the southern Gulf of St. Lawrence based on the September multispecies bottom trawl survey. DFO Can. Sci. Advis. Sec. Res. Doc. 2016/064. vii + 33 p.

TABLE OF CONTENTS

LIST OF TABLES IV
LIST OF FIGURES IV
ABSTRACT VI
RÉSUMÉ VII
1. INTRODUCTION 1
2. DATA 1
2.1 ABUNDANCE-AT-AGE 2
3. STATISTICAL MODEL 3
3.1 POOLED MODEL 4
3.2 AGE-DISAGGREGATED MODEL 5
4. RESULTS 6
4.1 POOLED MODEL 6
4.2 AGE-DISAGGREGATED MODEL 8
4.2.1 Catch-at-age estimates 10
5. DISCUSSION 10
5.1 BENEFITS OF THE APPROACH USED HERE 10
5.2 SMALL FISH COVERAGE 10
5.3 PELAGIC CATCHABILITY ISSUES 11
5.4 DIFFERENCES BETWEEN MODELS 11
5.5 FUTURE POSSIBILITIES 11
6. REFERENCES CITED 12
TABLES 14
FIGURES 16

LIST OF TABLES

Table 1. Number of otolith samples of the fall spawner component of Atlantic herring collected from catches in the September multi-species bottom trawl survey and used to derive the annual age-length keys.....14

Table 2. Pooled data model posterior summary statistics for the annual abundance index, standardized tows of 1.75nm, scaled to the CCGS Teleost, and with diel effects removed in numbers per tow. 15

LIST OF FIGURES

Figure 1. Stratification scheme used in the September multi-species bottom trawl survey. Coastal strata 401, 402 and 403 were added in 1984. 16

Figure 2. Pooled spatial distribution of herring (≥ 20cm) catches from the September multi-species bottom trawl survey, 1971 to 2014. Grey x's show zero catches while the area of red circles show catches on the log-scale. Catches were standardized for tow length. 17

Figure 3. Time series log-scale plot of observed catches of large herring (≥ 20 cm) in numbers. Catches were standardized for tow length. Observations were randomly slightly jittered on the horizontal scale so as to improve visibility. 18

Figure 4. Proportion of zero observations of large herring (≥ 20cm) in the September multi-species bottom trawl survey data set. 18

Figure 5. Mean number of herring per standardized tow versus water depth for all September multi-species bottom trawl survey catches 1971-2014. Grey bars show small fish (<20 cm) and black bars show larger fish (≥ 20 cm). 19

Figure 6. Log-scale empirical plot of standard deviation versus mean for each stratum by year. Catches were standardized for tow length. 19

Figure 7. Set of 20 B-spline basis functions used to model the diel effects. Linear combinations of these functions generate a periodic piecewise cubic function over the 24-hour daily cycle. For the age-disaggregated model 10 basis functions were used rather than 20. 20

Figure 8. Boxplots of posterior MCMC simulations of random effect error parameters for the pooled data model. 20

Figure 9. Log-scale boxplot of posterior year effects for the pooled data model. Boxes show the median (thick line) and quartiles while whiskers indicate 95% credibility intervals. 21

Figure 10. Log-scale boxplot of posterior stratum effects for the pooled data model. Boxes show the median (thick line) and quartiles while whiskers indicate 95% credibility intervals. 21

Figure 11. Bubble plot of log-scale year-stratum interactions for the pooled data model. Black and white circles indicate negative and positive deviations, respectively. 22

Figure 12. Estimated day-night (diel) effects as a function of the hour of day for the pooled data model. Shaded red bar show the range of sunrise and sunset times during the month of September. 23

Figure 13. Log-scale boxplots of MCMC simulated estimated vessel effects for the pooled data model. 24

Figure 14. Boxplot of posterior negative binomial precision parameters (α) estimated by stratum. 24

Figure 15. Boxplot of MCMC predicted means from the negative binomial model for standardized tows (1.75 nm tow, scaled to Teleost) with nuisance effects removed (vessel effects, diel effects, station effects). Red dots show the stratified mean estimates obtained using catch corrections in (Benoît 2006).....25

Figure 16. Boxplots of posterior MCMC simulations of random effect error parameters for year, stratum, age, diel effect, stratum by year interactions, year by age interactions, and diel effect by year interactions.....26

Figure 17. Log-scale boxplot of posterior year effects for the age-disaggregated model. Boxes indicate median, quartiles and whiskers indicate 95% credibility intervals.....26

Figure 18. Log-scale boxplot of posterior stratum effects for the age-disaggregated model. 27

Figure 19. Log-scale boxplot of age effects for the age-disaggregated model.....27

Figure 20. Estimated day-night (diel) effects as a function of the hour of day for the age-disaggregated data model. Shaded red bar show the range of sunrise and sunset times during the month of September.....28

Figure 21. Log-scale bubble plot of year by stratum interactions for the age-disaggregated model.....29

Figure 22. Log-scale bubble plot of diel by year interaction effects for the age-based model. Red bands show the range of sunrise and sunset times during September.....30

Figure 23. Bubble plot of year by age interaction terms, showing the annual catch-at-age deviations from the mean. Black circles indicate negative deviations whereas white circles indicate positive deviations. Circle area indicates the size of the deviation on the log-scale. Black and white diagonal bands show periods of low and high relative recruitment year classes, respectively.....31

Figure 24. Posterior stratified mean estimates of catch-at-age from the age-disaggregated model.....32

Figure 25. Scatterplots showing intra-cohort correlations by successive age groups based on the median the posterior predicted mean annual catch-at-age values. Units are in numbers per tow.....33

ABSTRACT

Fishery-independent abundance indices for the fall spawner component of Atlantic herring (*Clupea harengus*), one for all large fish (≥ 20 cm) and the other disaggregated by age for fall spawning fish only, are derived from the September multi-species bottom trawl survey in NAFO Div. 4T. Using a Bayesian approach, over-dispersion in the data is handled via a negative binomial error with log-linear mean and additive hierarchical priors over year, survey stratum, time of day, fish age and interaction terms. While indices remained highly variable, cohort tracking was apparent in the age- disaggregated index and inter-cohort correlations were moderate for ages 4 to 7 years. While uncertainty remains as to other types of systematic annual changes, there was little evidence of change in day-night catchability patterns of Atlantic herring over time.

Indices de l'abondance du hareng de l'Atlantique (*Clupea harengus*) du sud du golfe du Saint-Laurent tirés du relevé plurispécifique au chalut de fond réalisé en septembre

RÉSUMÉ

Deux indices d'abondance indépendants de la pêche applicables à la composante des reproducteurs d'automne de la population de hareng de l'Atlantique (*Clupea harengus*), un pour les poissons de grande taille (≥ 20 cm) et un pour les prises selon l'âge uniquement pour les poissons frayant à l'automne, ont été tirés des données du relevé plurispécifique au chalut de fond réalisé en septembre dans la division 4T de l'OPANO. Une approche bayésienne a été utilisée pour traiter la surdispersion des données au moyen de la distribution de l'erreur binomiale pour obtenir une moyenne définie par un modèle additif linéaire logarithmique et des valeurs *a priori* hiérarchisées selon l'année, la strate de relevé, la date et l'heure, l'âge du poisson et les interactions. La variation des indices était considérable, mais le suivi des cohortes était perceptible dans l'indice des prises selon l'âge, et les corrélations inter-cohortes étaient de niveau modéré pour les âges de 4 à 7 ans. Il y a toujours de l'incertitude quant aux autres types de variations annuelles, mais il y avait peu d'indices de changement des profils de capturabilité diurne et nocturne du hareng de l'Atlantique au fil du temps.

1. INTRODUCTION

Assessing southern Gulf of St. Lawrence (sGSL) Atlantic herring (*Clupea harengus*) presents a number of technical hurdles: it is a pelagic species which forms dense schools when feeding and spawning, performs annual migrations, and is composed of two separate stocks characterized by spring and fall spawning seasons. The schooling behaviour in particular leads to high variability in annual abundance estimates (i.e., large year effects) and potential bias.

A number of biomass and abundance indices have been developed or proposed for herring stocks in the sGSL. There are two fishery-dependent indices: a commercial gillnet CPUE index, and an experimental gillnet study over fall spawning grounds (Benoît et al. 2016; Surette et al. 2016). A fishery-independent index based on random stratified acoustic transect design (details in LeBlanc et al. 2015; see also LeBlanc and Dale 1996) has been produced from 1994 to 2014. However, this survey was found to track cohorts poorly, and is currently only used in stock assessments as an index of juveniles herring (i.e., ages 2 and 3). Fishery acoustic data collected by gillnet fishing vessels on spawning aggregations have been proposed to obtain spawning bed specific nightly biomass and exploitation rate indices (Claytor and Allard 2001; LeBlanc 2013; Surette et al. 2015) but these have not been used to date in the assessment models for the fall spawner component of Atlantic herring.

An annual September multispecies bottom trawl survey has been conducted since 1971 and provides some information on the distribution and abundance of many species in the southern Gulf of St. Lawrence. While originally designed to target groundfish, herring catches are common and this paper develops abundance indices for herring based on the bottom trawl survey catches.

2. DATA

The September multi-species bottom trawl survey has been performed annually since 1971. Sampling stations are distributed according to a stratified random design (Fig. 1) with strata defining areas of similar habitat and depth. The survey ranges in depth from 20 to 400 meters. Three in-shore strata were added in 1984; stratum 401 along the North coast of PEI, stratum 402 in the Northumberland Strait and stratum 403 in St. George's Bay (Fig. 1). The trawl was a Western IIA type with a liner in the cod-end with a mesh size of 19 mm. Further details on this survey may be found in Hurlbut and Clay (1990).

A total of five survey vessels have been used: the *E. E. Prince* (1971-1985), the *Lady Hammond* (1985-1992), the *CCGS Alfred Needler* (1992-2002, 2004-2005), the *CCGS Templeman* (2003) and the *CCGS Teleost* (2004-present). Comparative fishing experiments were undertaken for each change in survey vessel (Benoît and Swain 2003; Benoît 2006).

Herring catches for each set were weighed, and a subsample of the catch (up to 200 fish) was measured for length. Fish fork lengths were measured to a 1 cm precision from 1971 to 1996 and to a 0.5 cm precision from 1997 onward. Length-stratified otolith samples were collected for future ageing in the laboratory. Representative otolith age samples are only available from 1994 onward. Spawning group provenance (i.e., spring or fall spawners) was based on three methods. For juvenile herring of maturity stages 1 and 2, determination of the hatching season is based on size at capture and visual examination of otolith characteristics (Messieh 1972). Adult herring with ripe or spent gonads in September were assigned to the fall spawning component. Adult herring with non-ripe gonads were assigned their spawning component based on a gonadosomatic index (GSI) with an applied on a discriminant function (Cleary et al. 1982). The GSI is based on the length of the fish and its gonad weight (McQuinn 1989).

Two abundance index series were produced. The first is for herring larger than 20cm, spring and fall spawning components combined, for the period 1971 to 2014. This index was developed from a total of 6,378 survey sets over the time series. The second index is an age-disaggregated index using data from 1994 to 2011 for the fall spawning component only. This index includes data from a total of 3,226 sets over the time series. Repeated sampling at trawl stations occurred as part of vessel comparison or day-night experiments. These were taken into account in the analysis as station-level random effects. Auxiliary variables used in the analysis were the survey vessel, the survey year, the time of day of fishing, the stratum in which the tow was located, and the tow distance.

Catches are concentrated along the shores of northeastern New Brunswick, along the west, north and east shores of PEI, and in St. George's Bay up to Cheticamp in Cape Breton (Fig. 2). Herring are mainly found in coastal strata and mid-shore areas of deeper strata, such as strata 422, 429, 431, and 434 (Fig. 2). Few catches occur in the central and northeastern parts of the sGSL, though some small catches are sometimes caught in deep waters.

Nearly every year, there are catches which account for high proportions (up to 50%) of the total survey catches (Fig. 3). This indicates a high degree of over-dispersion in the count data, i.e., that herring have spatial distributions which are highly clustered, a consequence of its schooling and behaviour. Properly modelling the frequency of these large catches depends on distributional models sufficiently flexible to account for this variability. There are a large proportion of zero catches of herring across the time series (Fig. 4). The late seventies and early eighties saw a large proportion of zero catches, followed by a decline in the early eighties. From, that point there is no systematic trend over time with about 60% of sets having no herring in the sGSL (Fig. 4).

The vertical distribution of herring in the water column follows the upward migration of zooplankton at night while herring are more strongly associated with the sea bottom during the day. The contrast between daytime and nighttime catches are modelled and corrected for in the analyses.

The sGSL typically displays a high degree of water-column stratification during summer and fall and most species, including herring, have vertical distributions which vary as a function of water depth (Fig. 5). Small fish are primarily found in waters less than 30 meters while larger fish are found in the 30 to 60 meter range (Fig. 5). The depth profiles are sufficiently distinct that the overall distribution is strongly bimodal. As the bottom trawl survey targets waters 20 meters in depth or greater, it is possible that some portion of smaller fish may lie in coastal areas not covered by the survey, based on the pattern observed in Figure 5, although the bottom trawl survey appears to representatively sample the depth distribution of larger herring.

2.1 ABUNDANCE-AT-AGE

An age-disaggregated index for fall spawning herring was calculated from age samples gathered on the survey from 1994 to 2011. For sets with more than 200 herring, a length stratified sample of two fish per half-centimeter length interval was collected, frozen and processed in the laboratory. Data from these detailed samples included length, weight, sex, and otoliths were extracted for the determination of age. These samples were also used to determine the proportion of fall herring in herring catches. A maximum of two sets per stratum were sampled. A summary of the number of valid age samples per year is in Table 1.

For marginal lengths, i.e., very small or very large fish, the empirical age-length key is often deficient. For length categories with fewer than three otolith samples, the conditional probabilities were estimated using a multinomial probit regression model. The age frequencies were regressed against a set of 10 integrated B-spline basis functions defined over a set of

equally spaced knots over the range of 10 cm to 40 cm fish lengths. The coefficients of the basis functions were restricted to be positive so as to constrain the regression model monotonically increasing, as is expected for growth curves. To allow for heterogeneity in the errors, the variance of the underlying Gaussian distribution was defined as an exponential function of fish length. Maximum likelihood estimates for this model were obtained iteratively using the *optim* function from the R *stats* package (R Core Team 2013). For data deficient length categories (less than three samples), the predicted probabilities were used rather than the empirical proportions, as was the case for other length categories. By definition, this substitution was only sparsely applied, as it usually corresponds to rarely observed fish lengths, so that the overall contribution to the derivation of the age-disaggregated abundances is minimal. The conditional age probabilities of length categories with sufficient age samples were calculated as empirical probabilities. Age-length keys were calculated by 0.5 cm length category.

Only fall spawning herring were considered for the age-disaggregated analysis. The proportion of fall spawners was first calculated for each sampled stratum. The proportions were first calculated for each sampled set, and then a catch-weighted average of the values was produced for each stratum. For unsampled strata, the annual proportion of fall spawners was used, which was calculated as a catch-weighted average of the values from sampled strata. Annual proportions of fall spawners ranged from 0.73 to 0.91, with the last years (2005 to 2011) ranging from 0.83 and 0.88. Stratum-level proportions were then applied to herring length-frequency estimates for each set prior to the application of the age-length keys.

Catch-at-age was then calculated for each survey tow using the age-length key for the year applied to each tow's length-frequency. These count matrices by tow and age formed the inputs to the age-disaggregated analysis negative binomial model, used to generate the age-disaggregated abundance indices for 1994 to 2011.

3. STATISTICAL MODEL

Herring count observations for each catch were assumed to arise from a negative binomial distribution. The distribution is parameterized by its mean μ and a precision parameter α which controls the amount of dispersion in the data. The probability mass function is defined as:

$$\mathrm{P}[X=x] = \frac{\Gamma(x+\alpha)}{x!\Gamma(\alpha)}\left(\frac{\mu}{\mu+\alpha}\right)^{x}\left(\frac{\alpha}{\mu+\alpha}\right)^{\alpha} \quad (1)$$

The variance of this distribution is given by

$$Var[X] = \mu + \frac{\mu^2}{\alpha} \quad (2)$$

The negative binomial may be viewed as a type of over-dispersed Poisson distribution. Given a Poisson distribution with its mean conditioned on a gamma distribution, it can be shown that the resulting marginal count distribution is negative binomial (Johnson et al. 2005). As the precision parameter α becomes large, the negative binomial converges to the Poisson distribution and the variance converges to the mean μ as expected. For this application it is expected that there is considerable over-dispersion with respect to the Poisson, so the precision α is expected to be small.

The variance structure of the assumed negative binomial model is supported by empirical observations. An empirical log-scale scatterplot of the standard deviation versus the mean abundance of larger fish (≥ 20 cm) by year and stratum is shown in Figure 5. The piecewise linear character of the relationship between the two variables is a predicted consequence of the

quadratic mean-variance relation in equation (2). For low values of μ, the variance is dominated by the linear μ term (i.e. the distribution tends towards the Poisson), whereas as μ increases, the quadratic term dominates. On the log scale, this results in two linear domains with different slopes. To show this mathematically, we take the logarithm of the standard deviation from equation (2) which yields:

$$\ln\sigma = \ln\sqrt{Var[X]} = \frac{1}{2}\ln\left(\mu + \frac{\mu^2}{\alpha}\right) \tag{3}$$

Dividing on each side by $\ln\mu$ and taking the limit as $\mu \to 0$ we have, applying Hôpital's rule:

$$\lim_{\mu\to 0}\left(\frac{\ln\sigma}{\ln\mu}\right) = \frac{1}{2}\lim_{\mu\to 0}\frac{\ln\left(\mu + \frac{\mu^2}{\alpha}\right)}{\ln\mu} = \frac{1}{2}$$

Similarly, as μ becomes large:

$$\begin{aligned}
\lim_{\mu\to\infty}\left(\frac{\ln\sigma}{\ln\mu}\right) &= \frac{1}{2}\lim_{\mu\to\infty}\frac{\ln\left(\mu + \frac{\mu^2}{\alpha}\right)}{\ln\mu} \\
&= \frac{1}{2}\frac{\lim_{\mu\to\infty}\ln\left(\mu^2\left(\frac{1}{\mu}+\frac{1}{\alpha}\right)\right)}{\lim_{\mu\to\infty}\ln\mu} \\
&= \frac{1}{2}\frac{\ln\left(\lim_{\mu\to\infty}\mu^2 \times \lim_{\mu\to\infty}\left(\frac{1}{\mu}+\frac{1}{\alpha}\right)\right)}{\lim_{\mu\to\infty}\ln\mu} \\
&= \frac{1}{2}\frac{\ln\left(\lim_{\mu\to\infty}\mu^2 \times \frac{1}{\alpha}\right)}{\lim_{\mu\to\infty}\ln\mu} \\
&= \frac{1}{2}\lim_{\mu\to\infty}\frac{\ln\left(\mu^2 \times \frac{1}{\alpha}\right)}{\ln\mu} \\
&= \frac{1}{2}\lim_{\mu\to\infty}\frac{2\ln\mu - \ln\alpha}{\ln\mu} \\
&= \frac{2}{2} = 1
\end{aligned}$$

These limits show that for low values of μ, the slope of a logarithmic plot of the standard deviation versus the mean has a value of $\frac{1}{2}$, whereas as μ becomes large, the slope has a value of 1. This explains the piecewise linear pattern observed in the empirical plot in Figure 5.

Two regression models were developed; the first for a pooled herring index of fish larger than 20 cm, and a second age-disaggregated index. The goal of these analyses is to first predict the stratum means by year, then calculate stratified annual means from the estimates, using the stratum areas as weights.

3.1 POOLED MODEL

For the pooled index, the mean μ was defined be a log-linear function of year, set, stratum, time of day, vessel effect and tow distance:

$$\ln\mu_{ijklmn} = \ln\frac{d_{ijklmn}}{1.75} + \beta + \gamma_j + \delta_k + (\gamma\delta)_{jk} + \psi(h_l) + \upsilon_m + \varphi_n \tag{4}$$

where μ_{ijklmn} is the expected mean catch for tow i in year j, stratum k, time of day index l, vessel m at sampling station n. The model parameters are the intercept term β, the year effects

γ_j, the stratum effects δ_k, the year-stratum interaction term $(\gamma\delta)_{jk}$, the diel effect $\psi(h_l)$, the vessel effect υ_m and a sampling station level effect φ_n. An offset term was included to standardize the observed tow distance d_{ijklmn} to a standard distance of 1.75 nm. Diel (i.e. day-night) effects were modelled as a periodic piecewise cubic spline. A set of 20 cubic B-spline basis functions (Fig. 6), defined over equidistant knots over the 0 to 24 hour domain were evaluated at every half-hour of the day. The linear coefficients of these basis functions were given a hierarchical prior. The number of basis functions was chosen so as to only lightly smooth the data.

$$\psi(h) = \sum_{i=1}^{20} \theta_{\psi}^{i} B_{i,3}(h)$$
$$\theta_{\psi}^{i} \sim N(0, \sigma_{\psi}^{2})$$

The cubic B-spline functions are $B_{i,3}(h)$, where i in this case indexes the B-spline functions, h is the time of day, and θ_{ψ}^{i} are the associated B-spline coefficients. The following priors were placed on the remaining parameters:

$$\beta \sim N(0, 10^6)$$
$$\gamma_j \sim N(0, \sigma_{\gamma}^{2})$$
$$\delta_k \sim N(0, \sigma_{\delta}^{2})$$
$$\upsilon_m \sim N(0, \sigma_{\upsilon}^{2})$$
$$\varphi_n \sim N(0, \sigma_{\varphi}^{2})$$

A hierarchical log-normal prior over was placed over the dispersion parameter α to allow for the dispersion of the negative binomial distribution to vary within survey strata.

$$\alpha_j \sim LN(\mu_{\alpha}, \sigma_{\alpha}^{2})$$
$$\mu_{\alpha} \sim N(0, 10^3)$$

All standard error parameters were given half-Cauchy priors with scale parameter 10. Diffuse priors were used for all model parameters.

To calculate the mean abundance for the survey area, stratum means were first calculated from posterior simulations using the relation:

$$\ln \mu_{ijk} = \beta + \gamma_j + \delta_k + (\gamma\delta)_{jk} \qquad (5)$$

then summed over strata using areas as weights for each year.

3.2 AGE-DISAGGREGATED MODEL

Abundance-at-age indices were also produced using a negative binomial model. The log-linear equation (4) was modified to include contained three additional terms:

$$\ln \mu_{ijklmn} = \ln \frac{d_{ijklmn}}{1.75} + \beta + \gamma_j + \delta_k + (\gamma\delta)_{jk} + \psi(h_l) + \omega_a + (\gamma\omega)_{ja} + (\gamma\psi)_{jl} \qquad (6)$$

where $\omega_a \sim N(0, \sigma_{\alpha}^{2})$ is an additive age term indexed by age a, ranging from age 1 through 12, $(\gamma\omega)_{ja} \sim N(0, \sigma_{\gamma\omega}^{2})$ is a year by age interaction term, and $(\gamma\psi)_{jl} \sim N(0, \sigma_{\gamma\psi}^{2})$plus a diel by year interaction term to explore temporal variability in day-night catches.

A station effect was not included in the above model to speed up posterior model simulations, as the number of observations and effects were larger than for the pooled model. It was included in the pooled model as a way to include extra station-level variability for repeated sets.

In a different approach, the variability for regular and repeated sets was modelled by two different values for the precision parameter α. During the period considered for this age-disaggregated analysis (1994 to 2011), the only repeated sets were the comparative experiments between the *CCGS Needler* and the *CCGS Teleost* in 2004 and 2005. Only a single vessel effect was estimated for the *CCGS Needler* and *CCGS Templeman*. The *CCGS Teleost* was set as the reference vessel. The error parameters of the additional effects σ_a, $\sigma_{\gamma\omega}$ and $\sigma_{\gamma\psi}$ were given half-Cauchy priors with scale parameter 10.

To calculate catch-at-age across the survey area, stratum means by year and age were first calculated from posterior simulations using the relation:

$$\ln \mu_{ijk} = \beta + \gamma_j + \delta_k + (\gamma\delta)_{jk} + \omega_a + (\gamma\omega)_{ja} \quad (7)$$

then summed over strata using their areas as weights for each year.

Both models were fit using STAN (Stan Development Team 2016), a Bayesian MCMC sampling engine. Previous trials with OpenBUGS (Lunn et al. 2009) resulted in poor mixing of the posterior chains whereas STAN, which makes use of partial derivatives of the likelihood parameters in its samplers, results in much more efficient mixing of MCMC samples in cases where model parameters are highly correlated. For the pooled-data model, 10,000 MCMC simulations were generated after an initial burn-in of 500 values, whereas for the age-disaggregated model, only 1,000 simulations were generated given time limitations. Generated chains were homogeneous and had minimal auto-correlation.

4. RESULTS

4.1 POOLED MODEL

Boxplots of 1,000 MCMC posterior simulations for the pooled data model parameters and effects are shown in Figures 8 to 14.

Boxplots of error parameters, which control the scale of various model effects, are shown in Figure 8. Stratum effects had the highest mean variability of 3.3 reflecting a high degree of spatial variability of herring in the survey area. Station-level effects had the second largest mean variability at 1.71, followed by the diel effect at 1.06 and the year x stratum interaction effect at 1.06. The year effect, associated with overall interannual variability, had the smallest error at 0.81.

The initial period sees a decrease in year effects from 1971 to 1981, followed by a period of relatively high abundance from 1984 to 1995, followed by a decrease in 1996, then a relatively stable period at a slightly higher level from 1998 to 2004 (Fig. 9). The past few years saw a sharp increase in 2010 and 2011 followed by a decrease to 2014. The time series has a somewhat blocky character, as distinct abundance regimes appear, separated by discrete breaks at various points, such as in 1983 to 1984, 1995 to 1996 and more recently in 2009 to 2010.

Substantial variability occurs among strata, reflecting a wide range of mean herring abundances (Fig. 10). The strata surrounding PEI show the highest values, including the coastal strata 401 (north PEI) and 402, as well as the eastern stratum 421 and the western stratum 433. Strata along the northeastern coast of New Brunswick (419, 420 and 421) and well as St. George's Bay were also among the highest values (Fig. 10). Strata in the middle part of the sGSL, in the area around Bradelle Bank as well as along the Laurentian Channel had low values (Fig. 10).

The year by stratum interaction error term had a bimodal distribution, corresponding to two local solutions in the likelihood, the first with a relative error level of 0.2 with probability $\frac{1}{3}$, while the second solution was a mode at 1.5 occurring with probability $\frac{2}{3}$. The MCMC chains corresponding to these solutions were well mixed.

The year effects account for shifts in the spatial distribution of herring through time (Fig. 11). Among the more abundant strata, stratum 433 and 434 are seen to have more positive deviations starting in 1986, though large negative deviations were observed in 2014. Stratum 403 in St. George's Bay has increasing deviations from 2007 to 2012 and remained positive in 2013 and 2014. Stratum 422 had a series of low deviations starting in 1996 to 2009. Strata 419 and 420 had shorter periods of positive and negative fluctuations (Fig. 11). For less abundant strata, the period from 1971 to 1984 performed surveyed using the vessel *E.E. Prince* seems to have interaction effects that are generally of smaller magnitude than those of subsequent years. This suggests that some aspect of vessel catchability is not being properly accounted for, possibly due to size-based changes in trawl selectivity. Since the error parameter for the year by stratum interaction is bimodal, the effects themselves are generally bimodal, with a peak of values centered near zero plus another more diffuse off-centre peak.

Boxplots of diel effects for the pooled model are shown in Figure 12 along with sunrise and sunset times during the month of September. Fishing generally occurs at night when herring feed on vertically migrating zooplankton. In contrast, the bottom trawl used in the survey catches more herring at daytime during their inactive phase. Because the day is partitioned into half-hour periods, it is possible to see trends in relative catches as the day progresses. While there is some autocorrelation between the time periods, the 20 basis functions used are sufficiently flexible so that the trend is mainly data driven. Sunrise times in September range from 06:30 to 07:10 from the start to the end of the month which neatly matches a significant increase in diel effects corresponding to the 06:15-06:45 and 06:45-07:15 time periods. After this strong initial increase in catch levels, it is followed by a decrease of about 40% to a low at about 09:00, which gradually recovers as the time approaches 14:00. After this, there is a gradual decline in catches as the day approaches sunset at 19:00 to 19:50, at which point there is a sharp decline in catches up to 20:30. From then on the catchability remains low throughout the night hours though there may be some pulses of light changes in catch levels until morning, possibly related to changes in behaviour. On average catches during daytime were almost eight times as large as nighttime catches.

Boxplots of vessel effects are shown in Figure 13. Catches from past survey vessels, the *E. E. Prince*, the *CCGS Lady Hammond* and *CCGS Alfred Needler* were scaled to those of the current survey vessel, the *CCGS Teleost*. Catches for the *CCGS Templeman*, used only in 2003, were not scaled as there were no comparative experiments performed for this vessel. The mean fishing efficiency of the *E. E. Prince* was estimated to be 79% less than that of the *CCGS Teleost*, that of the *Lady Hammond* was found to be 36% less, and that of the *CCGS Alfred Needler* was found to be 22 % higher. We noted that the vessel effects for the *Lady Hammond* and *CCGS Alfred Needler* were not significantly different from that of the *CCGS Teleost*, based on 95% posterior credibility intervals.

The low values of the posterior precision parameters (i.e., α values) observed show that a high degree of over-dispersion is observed in all strata for large herring and that there is some heterogeneity between strata (Fig. 14). Small α values imply over-dispersion. Most values are smaller than one and more than half are smaller than 0.25. Values of α above 0.5 are associated with strata having low abundances, though stratum 403 in St. George's Bay is an exception.

Figure 15 shows annual stratified mean abundances, as calculated from the posterior stratum means, weighed by the stratum areas. Stratum means were calculated using equation (5) which uses only the intercept plus terms involving year and stratum effects. Diel terms, vessel effects, and station effects were treated as nuisance terms. In particular, marginalizing over diel effects imply that predicted stratum means represent average catches since the diel effects are zero-centered.

As was the case for year effects (Fig. 9), herring abundances have a somewhat blocky temporal pattern character with discrete phase changes appearing in the series. There is a decreasing trend in abundance from 211.9 fish per tow in 1971 to 7.9 fish per tow in 1981, followed by an abrupt increase from 22.1 in 1983 to 169.2 in 1984 which decreases to 46.3 in 1989, which rises sharply to 100.8 in 1990 and 110.7 in 1991. From a low of 17.2 fish per tow in 1996, the index rises to 50.8 fish per tow in 2002. The index decreases from 64.0 in 2005 to 26.1 in 2009, followed by another sharp increase to 122.0 in 2010, 74.7 in 2011 and 187.1 in 2012. The index then decreases abruptly to 74.1 in 2013 and 44.8 in 2014.

For comparison, empirical stratified means were calculated directly from data, adjusted for tow distance, vessel effects and day-level catches, using the method described in Benoît and Swain (2003) and Benoît (2006). Among the differences between the two methods, catches were adjusted to daytime catches rather than to an average value, and vessel effects had different values and no error model was assumed when calculating the stratified mean. Values between the two methods compare fairly well for the period 1984 to 1998, despite the difference in diel standardization, which as a rough estimate would imply that model-based values would be about 2.7 times higher if standardized to daytime levels. From 1999 onward, the empirical values are generally much higher than the model-based values. Calculated indices from this method for 2007, 2010 and 2012 were very high, being off the scale of the graph in Figure 15, owing to large year effects (values of 860.3, 934.5 and 3367.4, respectively).

4.2 AGE-DISAGGREGATED MODEL

Boxplots of 1,000 MCMC posterior simulations for the age-disaggregated model parameters and effects are shown in Figures 16 to 25. While the data for the age- disaggregated model were fall spawning herring of all sizes rather than all large herring for the pooled model, the overall effect patterns were broadly similar.

Boxplots of effect error parameters for the age- disaggregated model are shown in Figure 16. They are similar to those of the pooled model in Figure 8. The effect with the highest standard error parameter was stratum with a value of 4.3. Year effect was found to have the lowest variability with a value of 0.44. Age and day-night effects had similar levels of variability with values of 1.44 and 1.24, respectively. Interaction terms also had similar levels of variability between them, with mean standard error parameter values of 2.10 for stratum by year interactions, 1.09 for year by age interactions and 2.08 for diel by year interactions. In contrast to the pooled model, the stratum by year interaction had a unimodal posterior distribution rather than the skewed bimodal distribution obtained for the pooled model.

Little variability was observed between years with only a slight increasing trend from 2003 onward, though none of the effects seem to be significant (Fig. 17).

Stratum effects were more variable than those of the pooled model (Fig. 18). Strata 402 in the Northumberland Strait (Fig. 1) showed the largest decrease when the two models were compared, but this stratum is only partially sampled by the survey and the number of samples is limited. Stratum 417 and 418 showed increases in the age-disaggregated model, reflecting the presence of smaller fish in these strata catches.

Age 0 and age 11 fish are 27 times and 45 times less abundant than the most abundant age 4 fish, respectively (Fig. 19). While it is possible for younger fish to be less abundant than older fish in a given year if recruitment is sufficiently variable, the pattern observed here over the period 1994 to 2011 is probably a consequence of low selectivity of the trawl net. In addition, it was already noted that fish smaller than 20 cm were likely more coastally distributed (Fig. 5), which suggests that possibly important proportions of younger fish lie outside the survey bounds. The decreasing trend of age effects beyond age 4 (Fig. 19) reflects the cumulative effects of mortality across ages.\

The diel effect plot for the age-disaggregated model is shown in Figure 20. The observed trends are smoother than those observed for the pooled model in Figure 12, given that the number of underlying B-spline basis functions was 10 rather than 20, but the overall pattern was very similar. Smoothing also removed some of the variability observed during the daytime and nighttime periods in the pooled model. Estimates have a slightly higher variance owing to the more limited time period being considered relative to the pooled data set. Daytime catches were estimated to be about six times higher than nighttime catches.

Stratum by year interaction terms are shown in Figure 21. These show the time series of deviations from the average spatial distribution of herring over the study time period.

Year by diel interaction terms, the deviations of the average day-night catch pattern over time, are shown in Figure 22. While variable from year to year, the deviations seem to show no consistent change through time (Fig. 22). This indicates that relative catches between the day where herring interact with the bottom and night catches where midwater feeding occurs, does not seem to have changed during the study period. A hypothesis that herring interact more strongly with the sea bottom in recent years, particularly at night, is not consistent with the observed results. While the estimated diel effects are shown in half-hour increments, these are calculated from the weighted sum of 10 B-spline functions and 10 estimated coefficients, which explains the auto-correlated appearance of the series.

A bubble plot of year by age interaction effects on the logarithmic scale is shown in Figure 23. Year by age interaction terms show the deviation from the average catch-at-age through time. Diagonal bands of low and high recruitment pulses are readily apparent from the graph. Poor recruitment seems evident for the 1991, 1992, 1993, and 2003, year classes and to a lesser extent those of 1989, 2001, and 2002. Strong year classes are visible as high abundances of older fish from 1994 to 1998, whereas these largely disappear in subsequent years. Strong recruitment pulses are seen for the 1997 to 1999 year classes and more recently for the 2003 to 2006 yearclasses. Ages 3 to 7 in the middle part of the time series had small deviations, implying that the catch-at-age distribution for these ages resembled that of the global mean catch-at-age. Year 2005 had largely lower catches of younger fish and higher catches of older fish than those of adjacent years. Fish aged 0 seem to show more variability than other ages, probably due to the limited ability of the bottom trawl to catch these small fish or their unavailability to the trawl. Considering the degree of variability in observed herring catches, the presence of these cohort effects is encouraging and lends support that the RV survey may be capturing elements of the population dynamics, at least for fish aged 2 and older.

Precision α parameters had posterior means of 0.091 for regular sets and 0.153 for repeated sets. The larger value for repeated sets reflects the fact that repeated sets have more information, i.e., more than one count observation, than regular sets.

The vessel effect for the *CCGS Needler* and *CCGS Templeman* had a posterior mean of 0.86 with 95% credibility interval of 0.59 to 1.13, which corresponded to 236% more relative catches than the *CCGS Teleost*. This is a marked contrast with the estimate obtained in the pooled

model, which was 22% more than the *CCGS Teleost*. While the data sets are not identical, such a wide difference is surprising.

4.2.1 Catch-at-age estimates

Figure 24 shows a bubble plot of the stratified mean catch-at-age estimates in numbers per tow. Stratified values were calculated using equation (7). The start of the series shows a decrease in older fish from 1994 to 1997, then stabilizing in 1998 and 1999. The period from 2000 to 2005 shows a recruitment pulse through the indices, with a decrease in 2006. The period from 2008 to 2011 also sees a recruitment increase. The estimates for year 2005 were larger for ages 4, 5 and 6 than for adjacent years. This signal matches the large year by age interaction terms for large herring observed for 2005 in Figure 23.

Figure 25 shows bivariate plots of estimated abundances between fish of age a and age $a+1$ as well as linear correlations. The correlations are not particularly strong given the variability in the estimates. For use in a population model, correlations were felt to be sufficiently strong only for catches of herring age 4 ($\rho = 0.13$ or $\rho = 0.51$ when an outlying catch is removed), age 5 ($\rho = 0.49$ or $\rho = 0.36$ when the same outlying catch is removed) and age 6 ($\rho = 0.61$).

5. DISCUSSION

A fishery-independent pooled index and an age-disaggregated index for fall spawning Atlantic herring based on the catches from the September multi-species bottom trawl survey is presented. The concentrations of Atlantic herring schools are such that significant portions of the herring biomass reside in highly localized areas in the sGSL. While the stratified random survey sampling scheme yields an unbiased estimate of trawlable biomass, abundance estimates fluctuate markedly from year to year, depending on whether schools are encountered. The negative binomial distribution was chosen as a count model with a log-linear mean defined using hierarchical Gaussian priors over year, stratum, age, and diel effects, as well as interaction terms. The dispersion parameter was also given a hierarchical Gaussian prior over survey strata in the pooled model and a common value in the age-disaggregated model.

5.1 BENEFITS OF THE APPROACH USED HERE

The negative binomial was a reasonable choice to handle the considerable over-dispersion in these count data and the mean-variance relation assumed by the model was also deemed to be adequate. Hierarchical priors provide a simple way of modelling correlations between observations, such as temporal and spatial correlations, and aid in inferring group means where the data may be weakly informative. These correlations provide a degree of smoothing across group levels which control the influence of large catch observations, i.e., level means are shifted towards group means according to the amount of correlation. The cubic B-spline functions modelled the considerable variations in day-night catches and showed that estimated changes coincide with September sunrise and sunset times. While past investigations showed little evidence that September bottom trawl survey indices could track cohorts, cohort tracks are apparent in the year by age interaction terms and the predicted age-disaggregated indices. These lend support that the indices are able to track the dynamics of the herring population. Scatterplots between cohorts from adjacent years show reasonable correlations for fish aged 4 to 7, ages which could be used as indices within a population model.

5.2 SMALL FISH COVERAGE

In the age-disaggregated model, the abundance of small fish is probably underestimated, given the depth distribution of small versus large herring. A significant portion of smaller fish (< 20 cm)

composed of fish aged 0, 1 and 2 years, may be distributed in-shore of the 20-meter isobath which defines the bounds of the survey. Although 3-year old herring are seemingly well sampled by the survey, the low intra-cohort correlation at ages 3 to 4 years old suggests that samples may be unrepresentative. If the proportion of these small fish lying in unsampled shallower areas varies from year to year, their abundance indices would be biased.

5.3 PELAGIC CATCHABILITY ISSUES

A pelagic fish index based on bottom trawl sampling raises some issues. Herring catches represent a sizeable proportion, from 20% to 70%, of survey fish catches from the past 20 years. The question remains whether the estimates produced here are true indices. The requirement for a valid abundance index is that, for a given age or size, catches are a constant proportion of the population present at a given location. This may not be the case for Atlantic herring from this survey.

It has been suggested by McQuinn (2009) that the interaction of Atlantic herring with the bottom, and thus their availability to the trawl may have changed over time. For the sGSL herring data analysed in this document, the diel catchability is substantial however the inter-annual deviations from the mean diel pattern show little evidence of systematic change over the period 1994 to 2011. This means that the contrast between day and night catches does not seem to have consistently changed over time. The interaction of herring with the bottom changes with water depth and this may bring about concurrent changes in catchability. This factor was not examined in the current analysis, though a residual analysis might reveal patterns in catches which could be incorporated in future versions of the model. Finally, there are likely occasions when herring are caught incidentally as the trawl is being lowered or more probably, as it is being hauled. There is no way to account for this and for this analysis it was assumed that the catches occurred during the bottom trawling phase.

5.4 DIFFERENCES BETWEEN MODELS

While the pooled model and the age-disaggregated model had broadly similar results in terms of stratum effects (Figs. 10 and 18) and diel effects (Figs. 12 and 20), there were some notable differences. Estimated year effects from the pooled model from 1994 to 2011 were more variable than those from the age-disaggregated model which showed little variation and only a slight increasing trend in later years. The estimated vessel effects for the *CCGS Needler* differed markedly between the two models and the underlying reason is unclear. While the data analyzed were different in that only large fish were used in the pooled model and fall spawning fish in the age-disaggregated model, one might assume that the vessel effects would be similar, given that fall spawners comprise the majority (80% to 85%) of the herring caught during the survey.

There were some incoherences between the two analytical models. In addition to age-related effects, each had slightly different structure. The pooled model had a hierarchical dispersion parameter over survey strata rather than vessel-based estimates, plus a station effect. An interaction term for diel effect by year could also have been included in the pooled model as was the case in the age-disaggregated model.

5.5 FUTURE POSSIBILITIES

Formal model development could have been applied rather than the ad hoc model development method used in this paper. Such formal development would have included predictor variables and interaction terms based on demonstrable improvement in predictive performance or objective statistical criteria. In the Bayesian setting, such criteria include Bayes factors, the

commonly used Deviance Information Criterion (DIC) (Spielgelhalter et al. 2002) or alternate criteria as described by Watanabe (2010). In the model development in this paper, effects were added based on biological considerations and the scale of the estimated error parameters of the hierarchical priors. Our models may not adequately predict rare events such as the occurrence of large catches, thus inspection of the right-hand tails of the negative binomials could be critical in evaluating whether the frequency of these events is over- or under-predicted with respect to observations.

While cohort tracks are visible in age-disaggregated indices, the estimates remain highly variable with wide confidence intervals in predicted indices. Thus the indices developed in this paper provide only coarse information on changes in sGSL herring abundance. Options for other covariates or additional data could be considered in future analyses to reduce the amount of variability in the estimates.

In parallel with the September bottom trawl surveys, August mobile Sentinel surveys have been conducted by contracted fish harvesters following the same spatial experimental design as the September bottom trawl survey since 2003 (Savoie 2012). In these surveys, all Atlantic herring catches have been weighed and counted, and length-frequency samples have been recorded since 2007. The timing of these two surveys is sufficiently close that their data could potentially be combined to improve indices for years where they overlap. A similar approach has previously been applied to produce snow crab biomass indices by combining September bottom trawl survey and snow crab trawl survey data (Benoît and Cadigan 2014). As the Sentinel survey uses a larger-meshed trawl, size-based selectivity would need to be included as part of the age-disaggregated model.

The indices presented here could potentially be further stratified into stratum subgroups or sub-regions of the survey area, as is the case for the commercial CPUE indices of the fall herring fishery (Benoît et al. 2016).

Spring spawning herring could be similarly assessed using September survey data, with the caveat that age samples might need to be combined owing to the smaller proportion of spring spawning fish in the samples.

6. REFERENCES CITED

Benoit, H.P. 2006. Standardizing the southern Gulf of St. Lawrence bottom trawl survey time series: Results of the 2004-2005 comparative fishing experiments and other recommendations for the analysis of the survey data. DFO Can. Sci. Advis. Sec. Res. Doc. 2006/008.

Benoît, H.P., and Cadigan, N. 2014. Model-based estimation of commercial-sized snow crab (*Chionoecetes opilio*) abundance in the southern Gulf of St. Lawrence, 1980-2013, using data from two bottom trawl surveys. DFO Can. Sci. Advis. Sec. Res. Doc. 2014/082. v + 24 p.

Benoît, H.P., and Swain, D.P. 2003. Standardizing the southern Gulf of St. Lawrence bottom-trawl survey time series: adjusting for changes in research vessel, gear and survey protocol. Can. Tech. Rep. Fish. Aquat. Sci. 2505: iv + 95 p.

Benoît, H.P., LeBlanc, C., Surette, T., and Mallet, A. 2016. Background and data inputs for models evaluated as part of the 2015 southern Gulf of St. Lawrence fall-spawning herring assessment framework review. DFO Can. Sci. Advis. Sec. Res. Doc. 2016/063. ix + 50 p.

Claytor, R.R., and Allard, J. 2001. Properties of abundance indices obtained from acoustic data collected by inshore herring gillnet boats. Can. J. Fish. Aquat. Sci. 58: 2502 2512.

Crainiceanu, C.M., Ruppert, D, and Wand, M.P. 2005. Bayesian Analysis for Penalized Spline Regression Using WinBUGS. Journal of Statistical Software 14(14).

Gemlan, A., and Hill, J. 2009. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press. Cambridge, New York.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., and Rubin, D.B. 2013. Bayesian Data Analysis. Chapman & Hall/CRC Press, London, 3rd edition.

Johnson, N.L., Kemp, A.W., and Kotz, S. 2005. Univariate Discrete Distributions, 3rd edition. Wiley Series in Probability and Statistics. John Wiley and Sons, Inc.

LeBlanc, C.H. 2013. Nightly biomass estimates from acoustic data collected during the 2002 to 2012 herring gillnet fishing activities on fall spawning aggregations in the southern Gulf of St. Lawrence. Can. Tech. Rep. Fish. Aquat. Sci. 3040: vi + 20 p.

LeBlanc, C., and Dale, J. 1996. Distribution and acoustic backscatter of herring in NAFO divisions 4T and 4Vn, Sept. 23 - Oct. 08, 1995. DFO Atlantic Fisheries Res. Doc. 96/125. 28 p.

LeBlanc, C.H., Mallet, A., Surette, T., and Swain, D. 2015. Assessment of the NAFO Division 4T southern Gulf of St. Lawrence herring stocks in 2013. DFO Can. Sci. Advis. Sec. Res. Doc. 2015/025. vi + 142 p.

Lunn, D., Spiegelhalter, D., Thomas, A., and Best, N. 2009. "The BUGS project: Evolution, critique and future directions". Statistics in Medicine 28 (25): 3049–3067.

McQuinn, I.H. 2009. Pelagic fish outburst or suprabenthic habitat occupation: legacy of the Atlantic cod (*Gadus morhua*) collapse in eastern Canada. Can. J. Fish. Aquat. Sci. 66: 2256-2262.

Savoie, L. 2012. Results from the 2011 sentinel bottom-trawl survey in the southern Gulf of St. Lawrence and comparisons with previous 2003 to 2010 surveys. DFO Can. Sci. Advis. Sec. Res. Doc. 2012/068, iii + 56 p.

Spiegelhalter, D., Best, N., Carlin, B., and van der Linde, A., 2002. Bayesian measures of model complexity and fit (with discussion). J. Royal Stat. Soc. Series B 62(4): 583-639.

Stan Development Team. 2016. Stan Modeling Language: User's Guide and Reference Manual. Version 2.11.0. (accessed August 26, 2016).

Surette, T., LeBlanc, C.H., Claytor, R.R., and Loots, C. 2015. Using inshore fishery acoustic data on Atlantic herring (*Clupea harengus*) spawning aggregations to derive annual stock abundance indices. Fish. Res. 164: 266–277.

Watanabe, S. 2010. Asymptotic equivalence of Bayes cross-validation and widely applicable information criterion in singular learning theory. J. Machine Learn. Res. 11: 3571–3594.

TABLES

Table 1. Number of otolith samples of the fall spawner component of Atlantic herring collected from catches in the September multi-species bottom trawl survey and used to derive the annual age-length keys.

Year	Number of otolith samples
1994	244
1995	546
1996	314
1997	871
1998	826
1999	433
2000	447
2001	513
2002	407
2003	240
2004	578
2005	456
2006	341
2007	440
2008	265
2009	311
2010	355
2011	311

Table 2. Pooled data model posterior summary statistics for the annual abundance index, standardized tows of 1.75nm, scaled to the CCGS Teleost, and with diel effects removed in numbers per tow.

	Summary statistics					
Year	2.5%	25%	50%	Mean	75%	97.5%
1971	34.8	86.4	145.8	211.9	239.3	780.4
1972	20.3	48.9	77.8	102.4	120.7	339.2
1973	7.2	20.5	33.4	45.6	56.6	152.9
1974	8.6	20.3	34.4	47.4	58.0	150.5
1975	7.7	19.1	33.0	48.6	57.4	171.7
1976	11.3	26.9	43.0	58.1	71.1	181.4
1977	14.3	34.1	54.6	75.8	91.7	272.6
1978	4.8	13.4	24.8	37.3	45.8	150.3
1979	2.9	7.5	12.5	19.1	22.8	73.9
1980	2.7	7.6	12.6	18.0	20.6	59.4
1981	0.9	2.9	5.3	7.9	9.6	29.3
1982	2.9	7.8	13.0	17.9	21.7	58.7
1983	3.1	8.3	14.6	22.1	27.5	77.9
1984	28.1	73.2	121.2	169.2	203.9	593.6
1985	30.8	65.5	99.6	127.2	155.2	384.4
1986	14.8	35.5	56.7	78.0	93.2	268.4
1987	8.8	19.7	30.1	38.0	45.3	114.4
1988	10.1	25.2	41.4	54.6	66.6	194.1
1989	10.6	23.2	36.5	46.3	57.8	138.7
1990	22.6	51.6	77.4	100.8	123.7	304.7
1991	25.5	54.9	86.5	110.7	133.8	342.4
1992	11.7	23.0	33.8	39.7	50.5	96.6
1993	11.1	22.7	32.6	41.9	51.5	120.6
1994	14.8	32.1	49.3	60.1	71.7	175.2
1995	19.2	36.8	54.5	65.2	80.8	174.6
1996	5.1	9.6	13.9	17.2	20.7	48.2
1997	6.2	13.2	20.0	24.5	30.2	68.3
1998	7.9	16.1	23.7	27.7	34.7	75.9
1999	7.8	16.1	24.6	32.0	40.3	91.4
2000	11.5	24.7	37.1	47.0	56.8	142.6
2001	10.2	21.7	34.9	45.0	58.2	136.6
2002	9.8	23.9	39.4	50.8	63.8	167.4
2003	8.1	18.2	27.6	36.7	45.6	122.1
2004	8.3	16.3	22.6	27.5	33.3	78.5
2005	19.3	35.5	52.4	64.0	80.5	163.5
2006	5.6	10.7	15.8	21.8	25.4	78.7
2007	22.8	44.2	65.5	84.5	102.4	257.6
2008	13.2	26.1	37.1	45.8	55.8	124.0
2009	6.2	12.2	19.0	26.1	32.1	83.8
2010	35.0	63.6	96.5	122.0	146.4	321.7
2011	48.1	96.2	142.7	174.7	219.8	453.2
2012	32.1	83.0	145.8	187.1	235.6	545.7
2013	22.1	41.3	60.2	74.1	88.1	208.4
2014	13.0	25.0	35.8	44.8	55.3	120.5

FIGURES

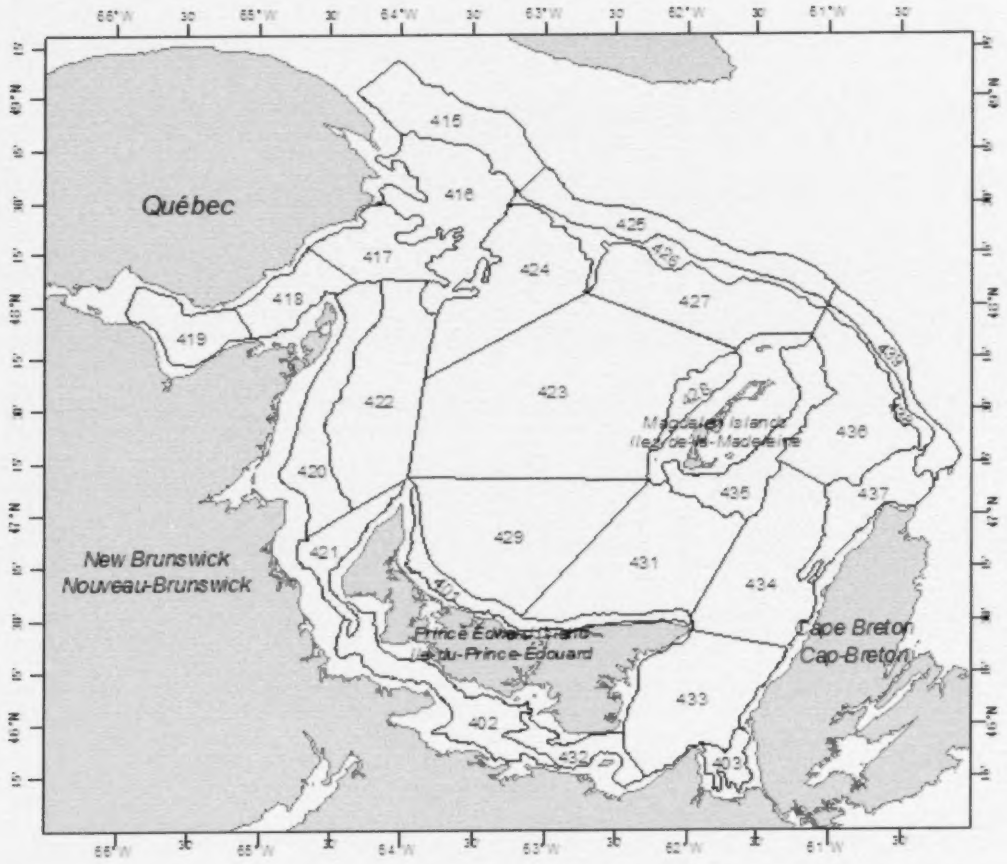

Figure 1. Stratification scheme used in the September multi-species bottom trawl survey. Coastal strata 401, 402 and 403 were added in 1984.

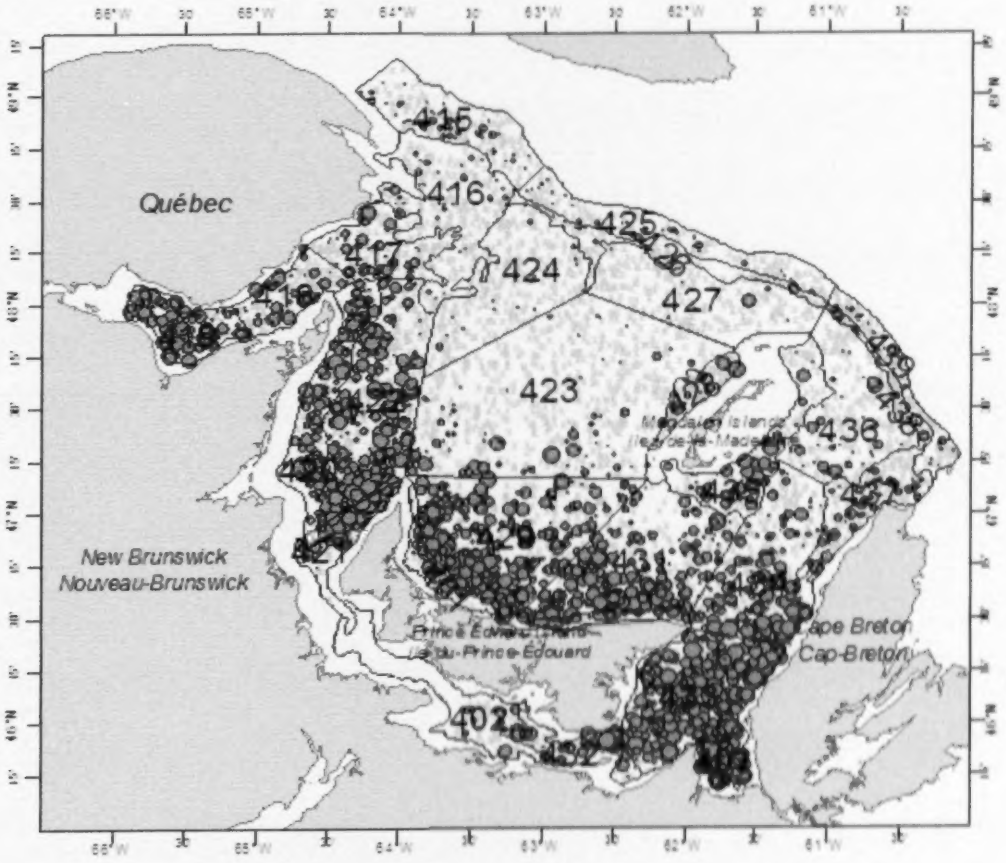

Figure 2. Pooled spatial distribution of herring (≥ 20cm) catches from the September multi-species bottom trawl survey, 1971 to 2014. Grey x's show zero catches while the area of red circles show catches on the log-scale. Catches were standardized for tow length.

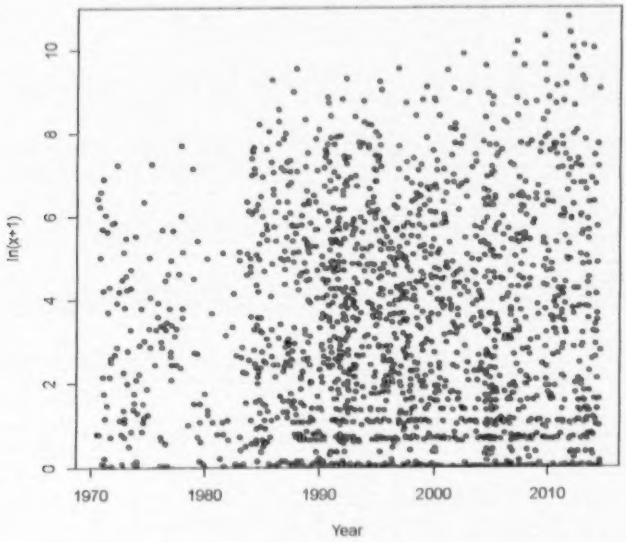

Figure 3. Time series log-scale plot of observed catches of large herring (≥ 20 cm) in numbers. Catches were standardized for tow length. Observations were randomly slightly jittered on the horizontal scale so as to improve visibility.

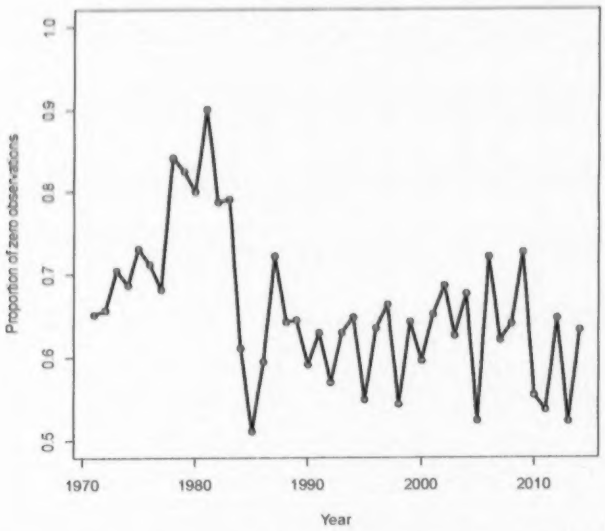

Figure 4. Proportion of zero observations of large herring (≥ 20cm) in the September multi-species bottom trawl survey data set.

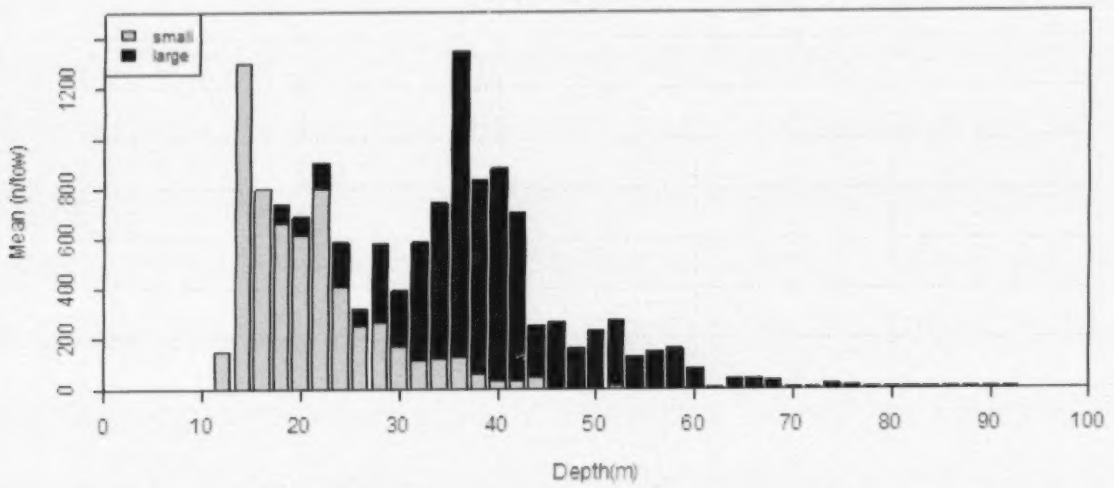

Figure 5. Mean number of herring per standardized tow versus water depth for all September multi-species bottom trawl survey catches 1971-2014. Grey bars show small fish (<20 cm) and black bars show larger fish (≥ 20 cm).

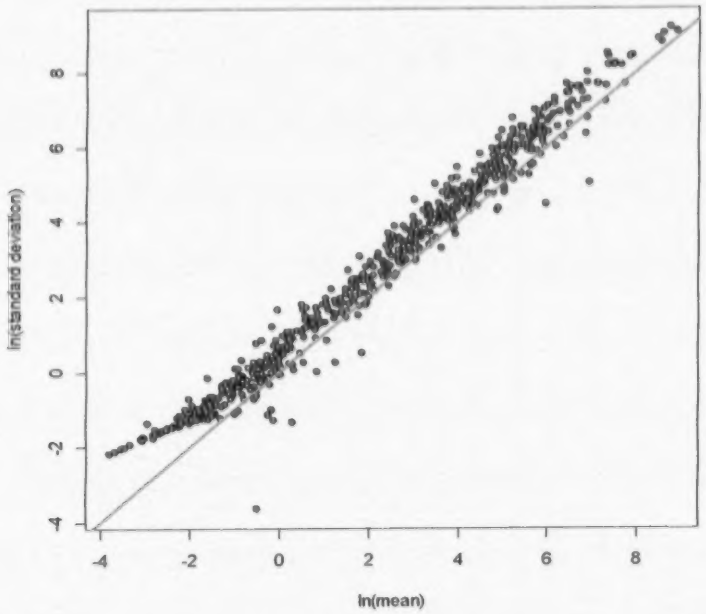

Figure 6. Log-scale empirical plot of standard deviation versus mean for each stratum by year. Catches were standardized for tow length.

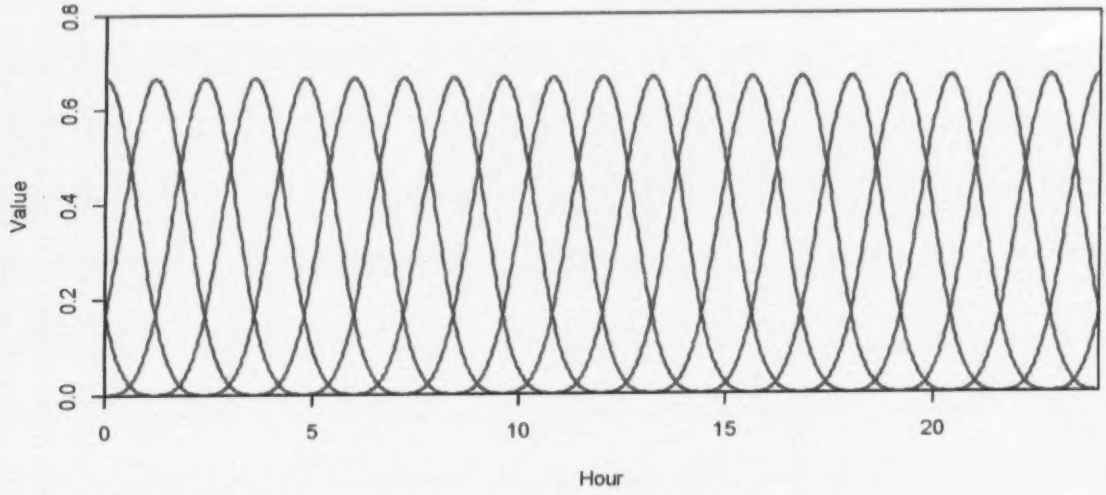

Figure 7. Set of 20 B-spline basis functions used to model the diel effects. Linear combinations of these functions generate a periodic piecewise cubic function over the 24-hour daily cycle. For the age-disaggregated model 10 basis functions were used rather than 20.

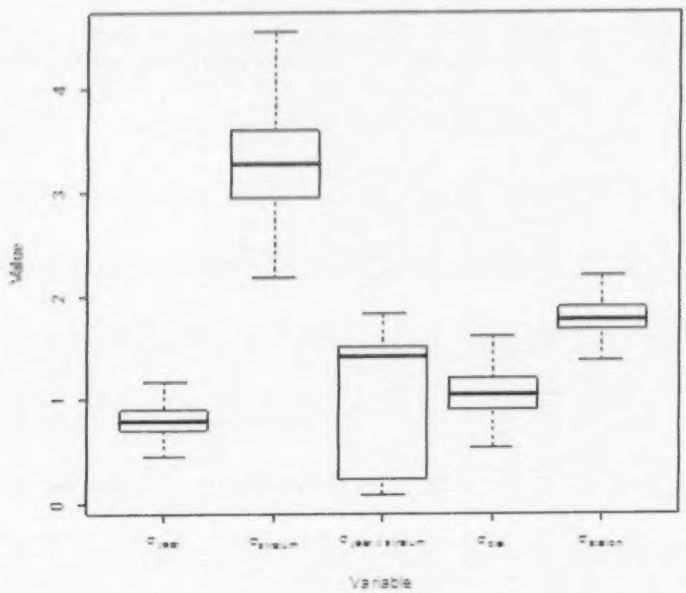

Figure 8. Boxplots of posterior MCMC simulations of random effect error parameters for the pooled data model.

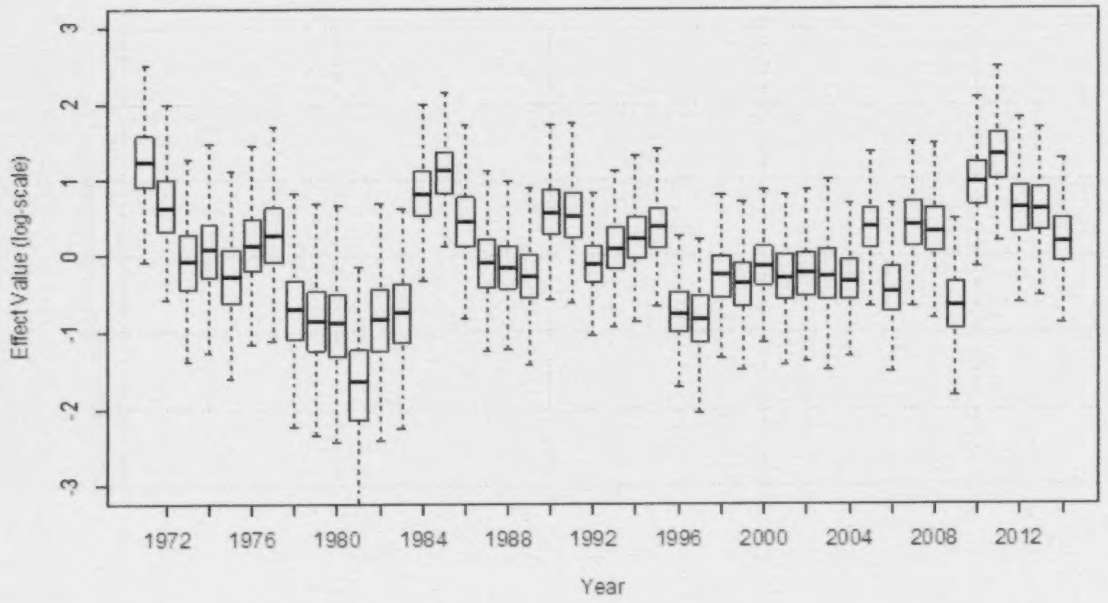

Figure 9. Log-scale boxplot of posterior year effects for the pooled data model. Boxes show the median (thick line) and quartiles while whiskers indicate 95% credibility intervals.

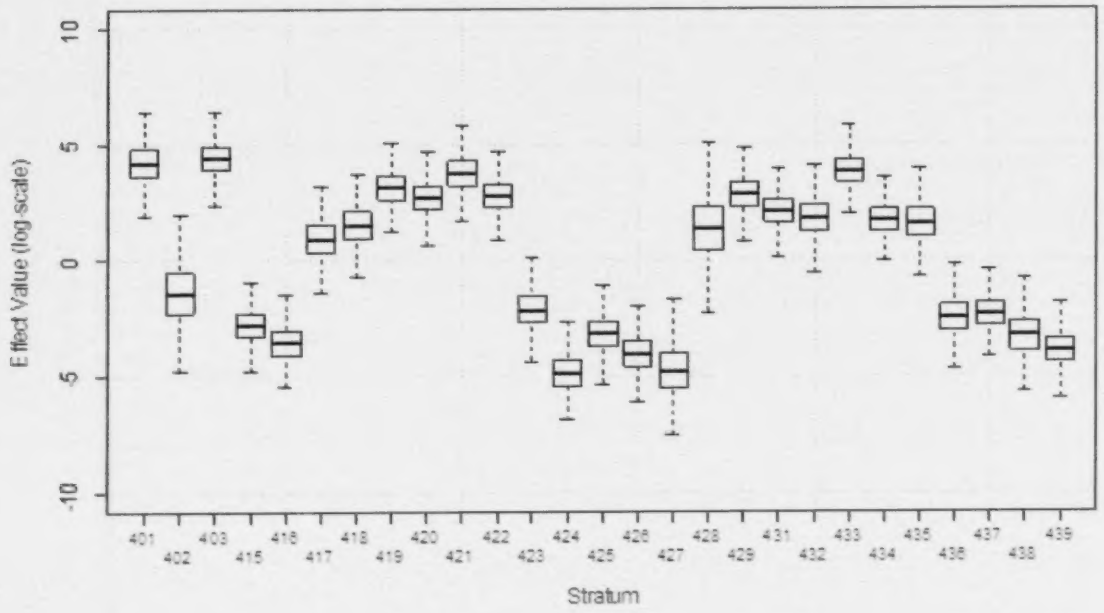

Figure 10. Log-scale boxplot of posterior stratum effects for the pooled data model. Boxes show the median (thick line) and quartiles while whiskers indicate 95% credibility intervals.

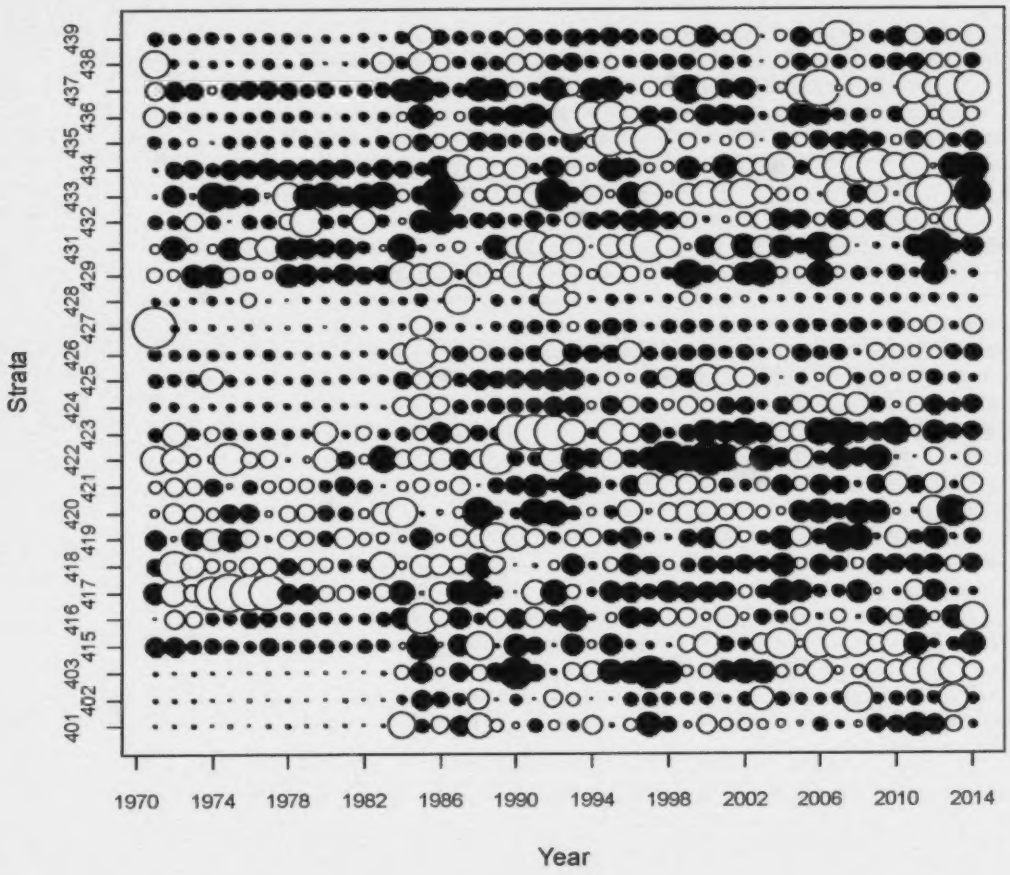

Figure 11. Bubble plot of log-scale year-stratum interactions for the pooled data model. Black and white circles indicate negative and positive deviations, respectively.

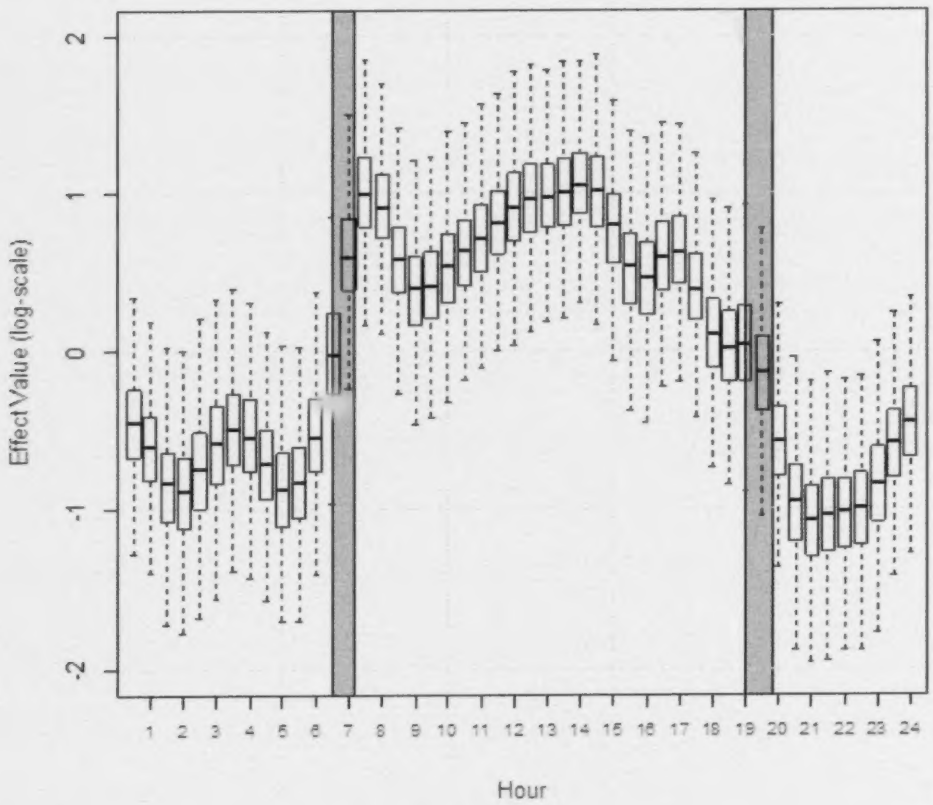

Figure 12. Estimated day-night (diel) effects as a function of the hour of day for the pooled data model. Shaded red bar show the range of sunrise and sunset times during the month of September.

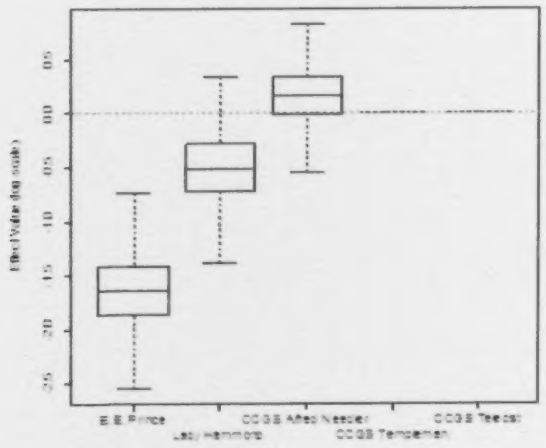

Figure 13. Log-scale boxplots of MCMC simulated estimated vessel effects for the pooled data model.

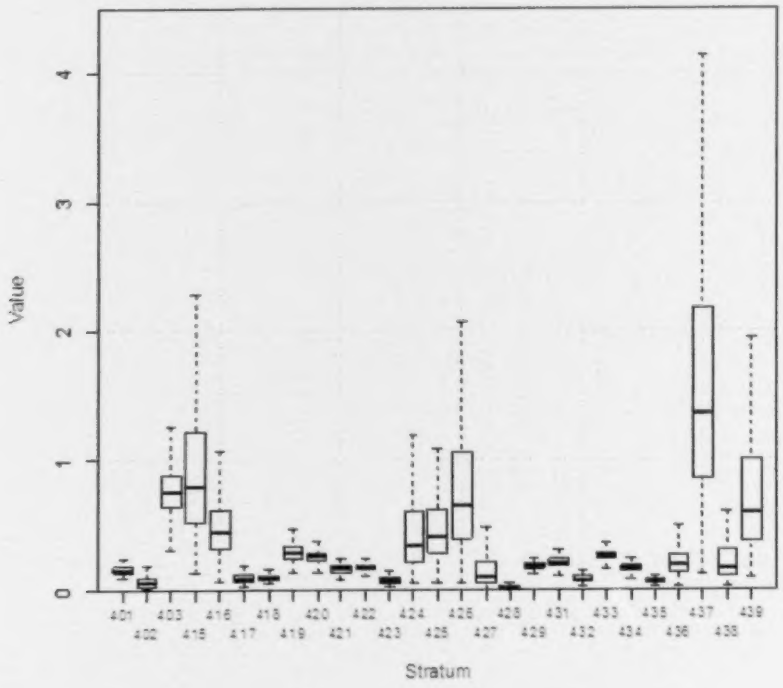

Figure 14. Boxplot of posterior negative binomial precision parameters (α) estimated by stratum.

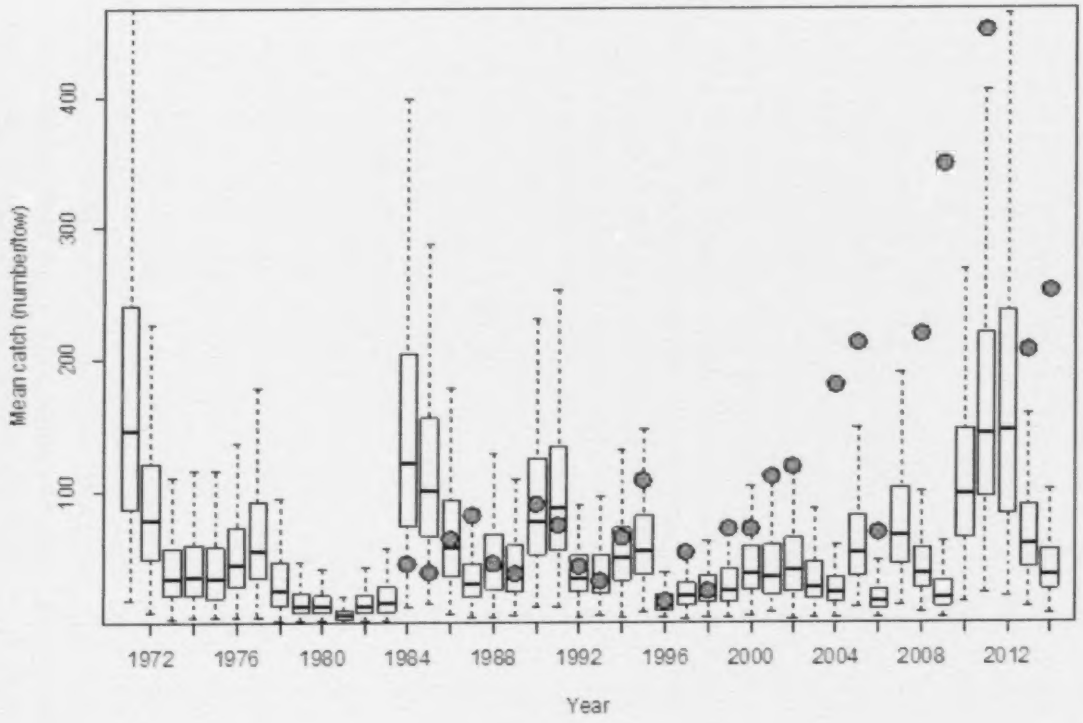

Figure 15. Boxplot of MCMC predicted means from the negative binomial model for standardized tows (1.75 nm tow, scaled to Teleost) with nuisance effects removed (vessel effects, diel effects, station effects). Red dots show the stratified mean estimates obtained using catch corrections in (Benoît 2006).

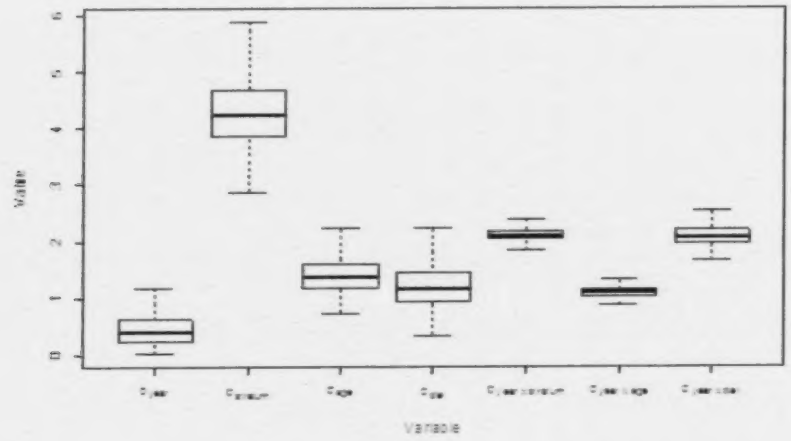

Figure 16. Boxplots of posterior MCMC simulations of random effect error parameters for year, stratum, age, diel effect, stratum by year interactions, year by age interactions, and diel effect by year interactions.

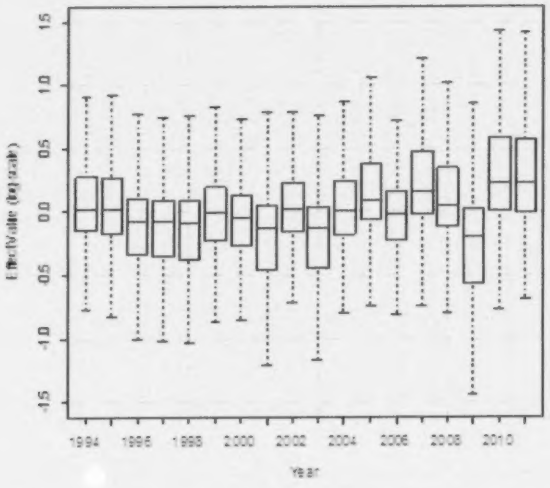

Figure 17. Log-scale boxplot of posterior year effects for the age-disaggregated model. Boxes indicate median, quartiles and whiskers indicate 95% credibility intervals.

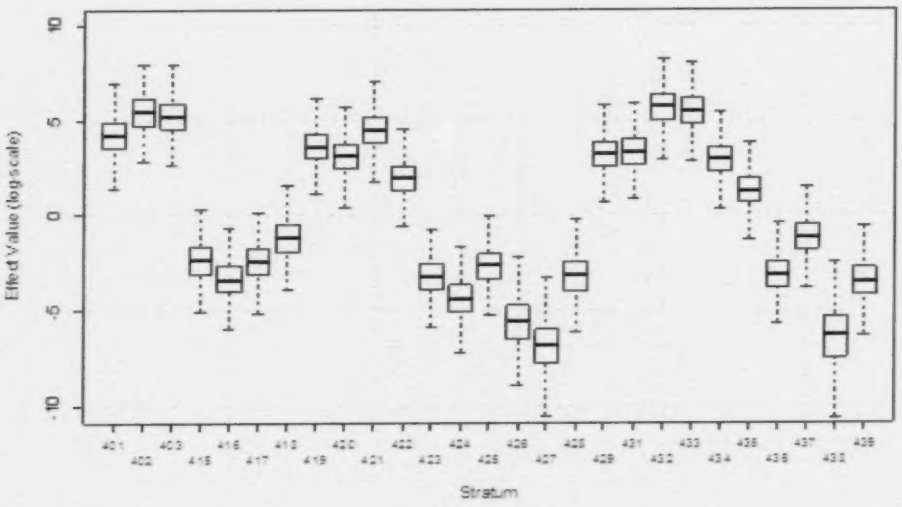

Figure 18. Log-scale boxplot of posterior stratum effects for the age-disaggregated model.

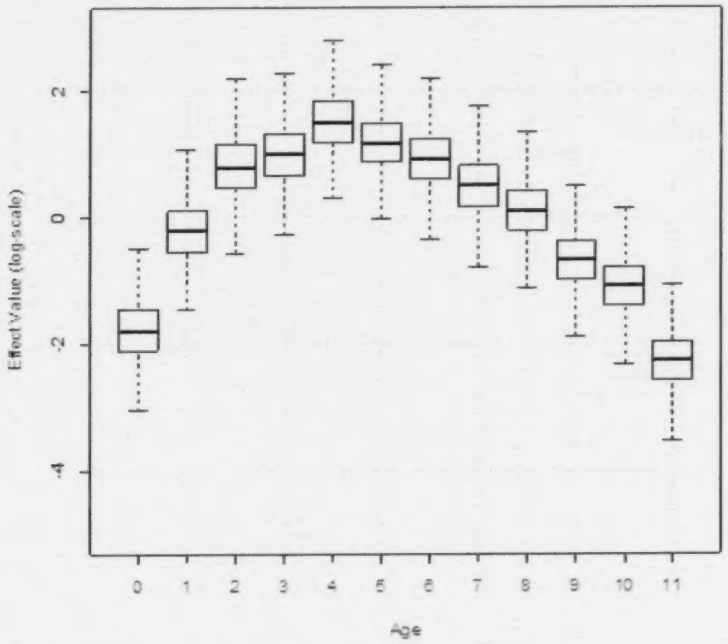

Figure 19. Log-scale boxplot of age effects for the age-disaggregated model.

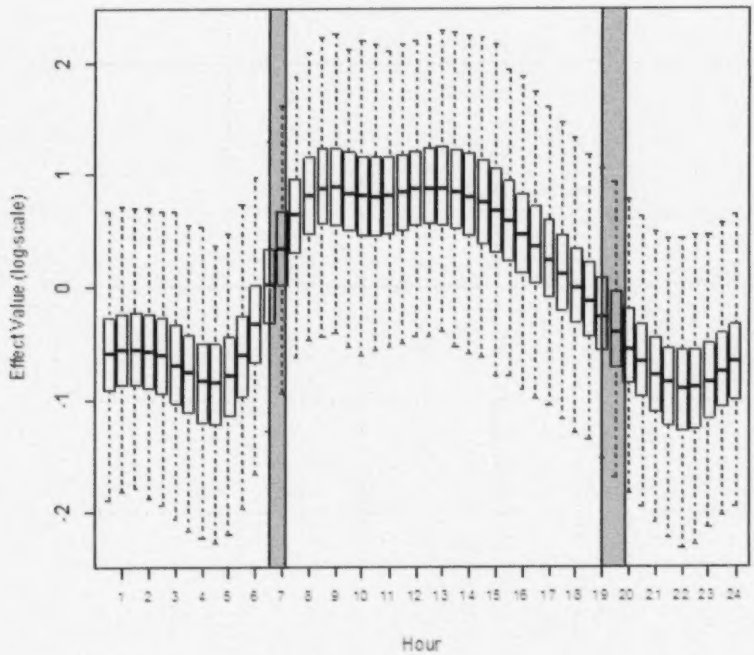

Figure 20. Estimated day-night (diel) effects as a function of the hour of day for the age-disaggregated data model. Shaded red bar show the range of sunrise and sunset times during the month of September.

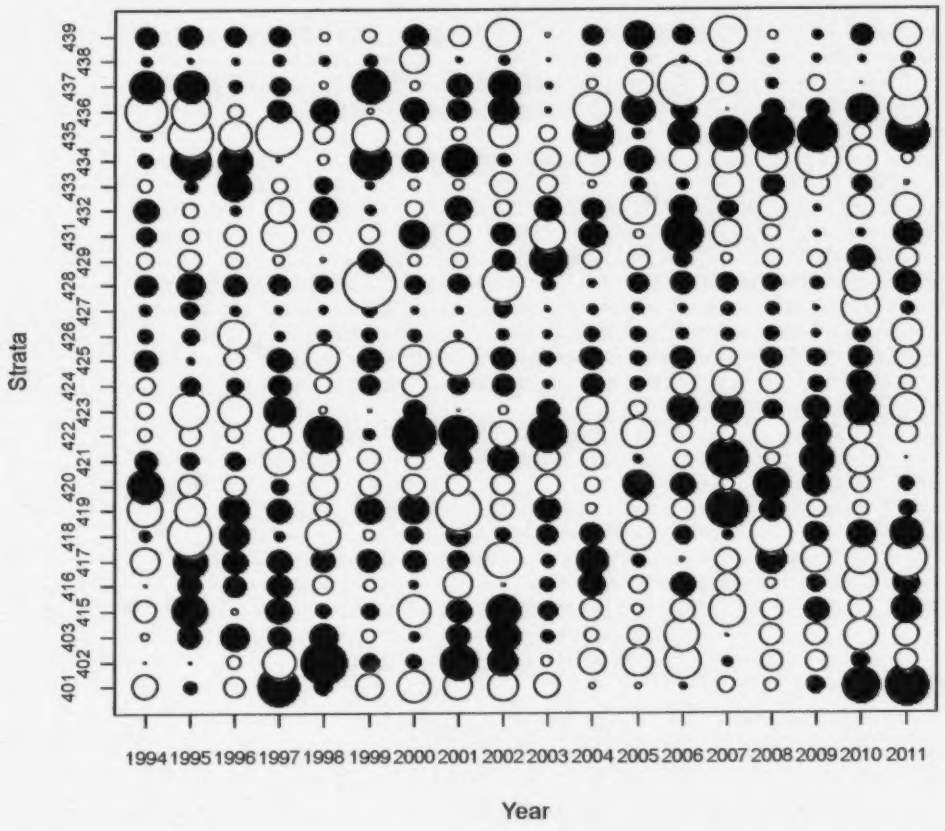

Figure 21. Log-scale bubble plot of year by stratum interactions for the age-disaggregated model.

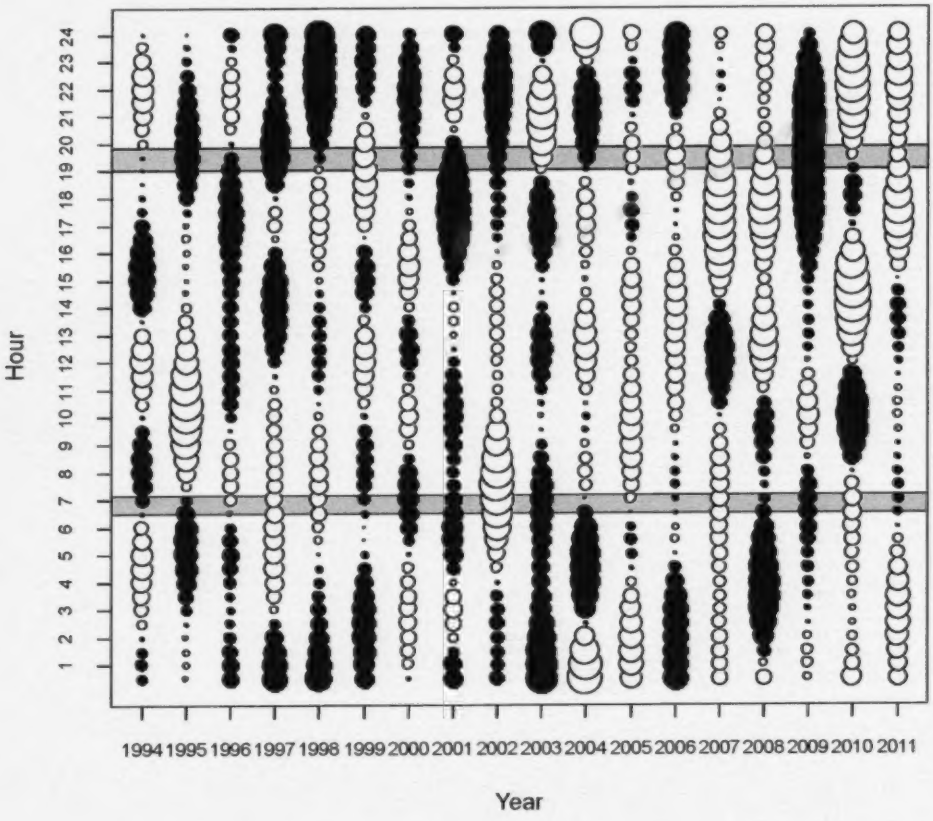

Figure 22. Log-scale bubble plot of diel by year interaction effects for the age-based model. Red bands show the range of sunrise and sunset times during September.

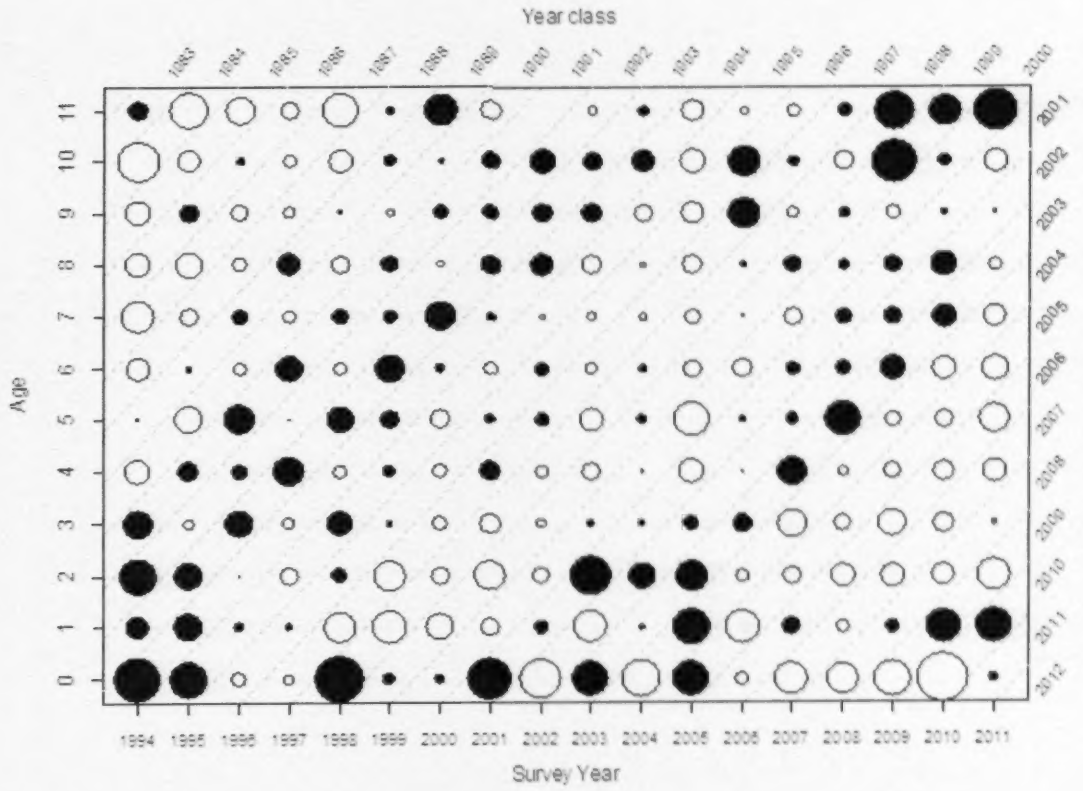

Figure 23. Bubble plot of year by age interaction terms, showing the annual catch-at-age deviations from the mean. Black circles indicate negative deviations whereas white circles indicate positive deviations. Circle area indicates the size of the deviation on the log-scale. Black and white diagonal bands show periods of low and high relative recruitment year classes, respectively.

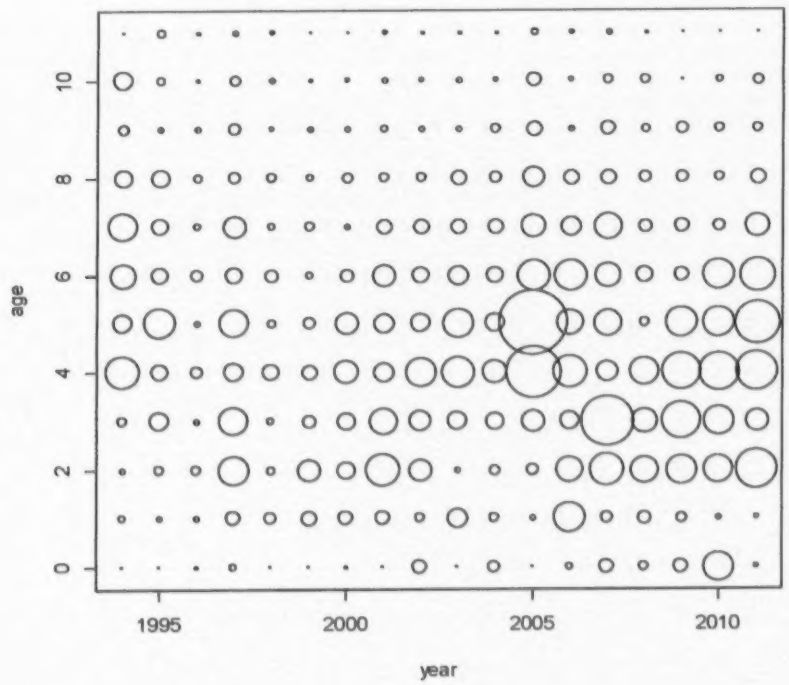

Figure 24. Posterior stratified mean estimates of catch-at-age from the age-disaggregated model.

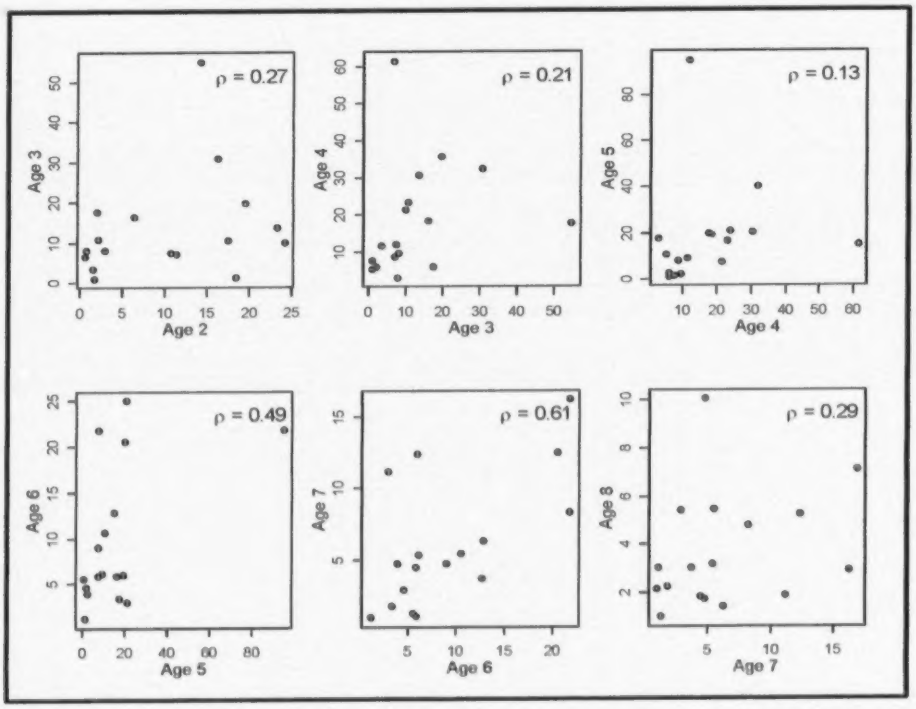

Figure 25. Scatterplots showing intra-cohort correlations by successive age groups based on the median the posterior predicted mean annual catch-at-age values. Units are in numbers per tow.